Method of enriching and/or separating prokaryotic DNA by means of a protein which specifically binds DNA containing non-methylated CPG motifs

The invention relates to a method of separating and/or enriching prokaryotic DNA or of
5   depleting said DNA from physiological liquids using a protein which specifically binds non-methylated cytidine-phosphate-guanosine dinucleotides (CpG motifs) of DNA, as well as to a kit for carrying out said method.

Infections caused by bacteria are one of the most frequent causes of inflammatory diseases.
10   For the prognosis of the clinical cause as well as, in particular, for timely selection of suitable therapeutic measures, early detection of the bacterial pathogens is of decisive importance.

In the detection of bacterial pathogens use is made even today, above all, of different methods of cultivating cells. However, current studies clearly show the poor suitability of culture-
15   dependent methods for detection of pathogens (Hellebrand W., König-Bruhns C., Hass W., Studie zur Blutkulturdiagnostik im Jahr 2002, Poster Jahrestagung der Deutschen Gesellschaft für Hygiene und Mikrobiologie, Göttingen 2004; Straube E (2003) Sepsis – microbiological diagnosis. Infection 31:284). According to these studies, it was possible to determine pathogens in only approximately 15-16 % of all blood cultures examined. As a result of the disadvantages
20   of these methods, increased efforts were made to find alternatives, especially during the past decade, simultaneously with the rapid technological development in molecular biology. First reports on the use of culture-independent methods of detecting bacterial pathogens, based on the principal of the polymerase chain reaction (PCR), date back to the early 1990s. Thus, for instance, Miller and colleagues (Miller N J Clin Microbiol. 1994 (Feb;32(2):393-7) were able to
25   show that culture-independent methods are superior to the classic techniques of cultivation and microscopy for detection of *mycobacterium tuberculosis*. However, further molecular-biological methods based on the detection of pathogen-specific nucleic acids have gained importance (e.g. M. Grijalva et al. Heart 89 (2003) 263-268; Uyttendaele M et al. Lett Appl Microbiol. 2003;37(5):386-91; Saukkoriipi A et al. Mol Diagn. 2003 Mar;7(1):9-15; Tzanakaki G et al.
30   FEMS Immunol Med Microbiol. 2003 Oct 24;39(1):31-6).

In addition to the high specificity of such molecular-biological methods, the reduced time expenditure is to be mentioned as a substantial advantage over conventional culture-dependent methods. However, the sensitivity of the detection of prokaryotic DNA directly from body fluids and not from pre-treated testing material as compared to the culture of microorganisms has been much too low so far. An amount of nucleic acids of bacteria sufficient for the directed detection of pathogens from testing material, which is not pre-treated, is achieved to a limited extent only also with respect to the 16S-rRNA analysis, by means of PCR of the 16S region on the bacterial chromosome and the subsequent sequence analysis of the PCR fragment, because in most cases, several copies of the segment encoding 16S-rRNA are found on the chromosome. The direct specific detection of pathogens by means of 16S-rRNA analysis requires that only one species of pathogen is present in the sample to be examined. If there are different species of pathogens in the sample, specific detection by sequencing of the 16S-rRNA region is not possible, because the primers used are universal for most bacteria. Further, it is a prerequisite to the detection of pathogens by 16S-rRNA analysis that the bacteria to be detected are present in the metabolic phase and sufficiently express 16S-rRNA. This is usually not the case, in particular in patients subject to calculated antibiotic therapy. Moreover, expression of certain pathogenicity factors of bacteria does not occur at all times, although the corresponding genes are present in the bacterial genome. As a result, erroneously negative results are transmitted to the clinical physician. Thus, selective antibiotic therapy may be initiated either not at all or much too late. In such cases, the physician has to rely on his knowledge gained by experience and on general guidelines (such as those of the Paul Ehrlich Foundation) and will therefore effect a much too general antibiotic treatment. The unspecific use of antibiotics bears a number of risks, not only for the individual patient (such as unnecessary side effects in the form of renal damage etc.), but also for the entire society (e.g. the development of additional antibiotic resistances, such as MRSA (methicilline-resistant *Staphylococcus aureus*, etc.). Therefore, the detection of clinically meaningful pathogenicity factors and resistances of bacteria at the chromosomal level and at the plasmid level, i.e. ultimately on the DNA level, provides considerable advantages for the diagnosis of many infectious diseases but also of sepsis. This applies even more because, at this level, a distinction can also be made between pathogenic and commensal bacteria.

Most frequently, the detection of pathogen-specific nucleic acids is effected by nucleic acid amplification techniques (NAT), such as the amplification of the prokaryotic DNA by means of the polymerase chain reaction (PCR) or the ligase chain reaction (LCR), respectively. The high specificity and fast availability of the results is contrasted by the susceptibility to interference by contamination or by strongly reaction-inhibiting factors of clinical samples.

In a conventional PCR detection method, successful detection of pathogens in the blood theoretically requires at least 1 target DNA of the pathogen to be present in 10 μl of blood. This corresponds to approximately 100 targets in 1 ml of blood or 1,000 targets in 10 ml of blood, respectively. Things are different with regard to the blood culture for detection of infection
5   pathogens. In this case, the lower detection limit is approximately 3-5 bacteria per 10 ml of blood.

This detection limit is presently not reached yet by PCR methods, not even by those which have their target sequences in the vicinity of the 16S-rRNA region on the chromosome. Although
10   several regions encoding 16S-rRNA are located on the bacterial chromosome, in most cases 3 to 6, the prerequisite that at least one molecule of the template DNA is located in the PCR reaction mixture is not met.

Improved diagnostic safety is to be expected of PCR methods whose specific target sequences
15   encode species-specific proteins, either in the chromosome or on plasmids of the microorganisms. The above remarks with respect to the detection limit also apply here. Especially under the action of a current antibiotic therapy, growth of the pathogens can be considerably decelerated, limited or blocked, even if the antibiotic employed ultimately does not have an optimal effect. This situation is often found especially in patients who are already
20   receiving antibiotic treatment and in whom, therefore, no disease-causing bacteria can be grown from blood cultures or other samples (such as for example tracheal smears, broncho-alveolar lavages (BAL) etc.).

Due to insufficient sensitivity, the detection of pathogen-specific nucleic acids without an
25   amplification step by direct detection of prokaryotic DNA (probe technique, FISH technique) is of diagnostic importance only at a sufficiently high germ count in the test material.

The essential problems of the detection of prokaryotic DNA for identification of bacterial pathogens in body fluids consist, in addition to PCR-inhibiting ingredients in the test material,
30   mainly in the low concentration of prokaryotic DNA and the resulting excess of eukaryotic DNA versus prokaryotic DNA. In this connection, in particular, competitive processes in DNA analysis as well as the quantity of prokaryotic DNA can be regarded as a hindrance to qualitative and quantitative detection of pathogens.

35   The usual methods of DNA isolation enrich the total DNA of a body fluid so that the ratio of host DNA to microbial DNA may be between $1:10^{-6}$ and $1:10^{-8}$. This difference makes the difficulty in detecting microbial DNA in body fluids quite easy to understand.

Prokaryotic DNA differs from eukaryotic DNA, for example, by the presence of non-methylated CpG motifs (Hartmann G et al., Deutsches Ärzteblatt, Jg. 98/15:A981-A985 (2001). In the prokaryotic DNA, 16 times more CpG motifs are present than in eukaryotic DNA, which contains such motifs only temporarily, for example in cancer cells or promoter regions. These motifs are

5 not methylated in prokaryotic DNA, whereas the majority of them are methylated in eukaryotic DNA, which further augments their distinctiveness. Non-methylated CpG motifs are non-methylated desoxycytidylate-desoxyguanylate-dinucleotides within the prokaryotic genome or within fragments thereof.

10 It is further known that diagnostic statements for cancers can be derived from different methylation patterns within the human DNA (Epigenetics in Cancer Prevention: Early Detection and Risk Assessment (Annals of the New York Academy of Sciences, Vol 983) Editor: Mukesh Verma ISBN 1-57331-431-5). Methylated and non-methylated cytosines in the genome allow tissue-specific but also disease-specific patterns to be identified. The specific methylation

15 patterns of a disease allow, on the one hand, diagnosis at a very early point in time and, on the other hand, molecular classification of a disease and the likely response of a patient to a certain treatment. For detailed information on this, see, for example, Beck S, Olek A, Walter J.: From genomics to epigenomics: a loftier view of life.", Nature Biotechnology 1999 Dec;17(12):1144, on the homepage of Epigenomics AG (http://www.epigenomics.de), or WO 200467775.

20

Cross at el. showed that it is possible to separate differently methylated genomic human DNA by binding the methylated CpG motifs to a protein (Cross SH, Charlton JA, Nan X, Bird AP, Purification of CpG islands using a methylated DNA binding column, Nat Genet. 1994 Mar;6(3):236-44). Thus, this method serves to bind DNA containing methylated CpG motifs.

25 Sufficient isolation of non-methylated and methylated DNA is not possible for technical reasons, because the protein used also weakly binds non-methylated DNA. It is also not possible with these methods to enrich non-methylated DNA, because the capacity of the protein used is not sufficient to separate non-methylated DNA to a sufficient extent in the case of a high excess of methylated DNA. Further, due to the binding of the methylated DNA, the initial volume in which

30 the non-methylated DNA is present, remains unchanged so that no enrichment is achieved.

Thus it would be desirable to separate non-methylated DNA from methylated DNA and to be able to enrich non-methylated DNA so as to separate prokaryotic DNA from eukaryotic DNA or differently methylated human DNA, respectively, from each other. In addition, it would be

35 desirable and of great interest in terms of health economics if the isolation and enrichment of non-methylated DNA could also be obtained from a mixture (for example, full blood) which is characterized by a great excess of methylated DNA.

It is known from Voo et al. that human CpG-binding protein (hCGBP) is capable of binding non-methylated CpG motifs. This publication describes the transcription-activating factor hCGBP which has been shown to play a role in the regulation of gene expression in CpG motifs.

5 EP 02020904 shows a method which enables isolation and enrichment of prokaryotic DNA from a mixture of prokaryotic and eukaryotic DNA by binding the prokaryotic DNA to a protein which specifically binds non-methylated DNA.

Therefore, it is an object of the present invention to provide a method of separating and/or

10 enriching prokaryotic DNA from examination samples having a high content of eukaryotic DNA, in particular from patients with infections.

According to the invention, this is achieved by a protein binding non-methylated CpG motifs, said protein having a 25 % to 35 % homology, in particular approximately 27.6 % homology,

15 with wild type CPGB protein and is shortened with respect to the latter, to the length of the binding site at maximum.

The human CPGB protein (cf. Voo et al., Mol Cell Biol. 2000 Mar; 20(6): 2108-21) is referred to hereinafter as wild type CPGB protein (or CPGbP656). The protein according to the invention is

20 referred to hereinafter as CPGbP181. The protein described in EP 02020904, which is a shortened variant of the wild type CPGB protein and served as the basis for the protein according to the invention, is referred to hereinafter as CPGbP241.

The invention will be described below with respect to the Figures, wherein:

25

Figure 1   shows the amino acid sequence of CPGbP181 (in bold print) compared with the wild type CPGB protein (CPGbP656) and CPGbP241 (printed in italics);

Figure 2   shows the DNA sequence and translation to the amino acid sequence of the

30   complete CPG-binding protein CPGbP656, wherein the shortened CPG-binding peptides CPGbP241 (bold) and CPGbP181 (italics) are shown;

Figure 3   shows a PCR of streptococci-DNA in human blood;

35 Figure 4   shows a nested PCR with the PCR products from the primary PCR approach of Figure 3 as template;

Figure 5   shows a gel retardation experiment;

Figure 6   shows a further gel-retarding experiment;

Figure 7   shows the elution of calf thymus DNA and pUC18emm by rCpG-181 sepharose,

5   and

Figure 8   shows the determination of the eluted DNA in the fractions by measurement of the extinction at 254 nm as a function of the NaCl gradient;

10 Figure 9   shows the results of PCR after enrichment of prokaryotic DNA from a DNA mixture of *staphylococcus aureus* and human DNA using coupled CpGbP-181 protein on CNBr sepharose, and

Figure 10   shows results of PCR after enrichment of prokaryotic DNA from a DNA mixture

15   of *staphylococcus aureus* and human DNA using coupled CpGbP-181 protein on AH-sepharose.

The wild type CPGB protein CPGbP656 binds non-methylated CpG motifs of prokaryotic DNA, thus forming a protein-DNA complex. This complex may be or become attached to, for example,

20 a carrier, whereby separation and/or enrichment of DNA can be effected. The present invention is now based on the surprising finding that a protein which is shortened relative to the wild type CPGB protein (CPGbP656 comprising 656 amino acids) and presenting 25 % to 35 %, in particular approximately 27.6 %, homology with the wild type CPGB protein, has improved binding properties over non-methylated CpG motifs of prokaryotic DNA than the wild type CPGB

25 protein and variants thereof with a homology of 80% or more. An example of such shortened protein is CPGbP181 with 181 amino acids.

Prokaryotic DNA differs from eukaryotic DNA, for example, by the presence of non-methylated CpG motifs (Deutsches Ärzteblatt, Jg. 98/15: A981-A985 (2001)). The invention is based on the

30 finding that eukaryotic DNA and prokaryotic DNA differ in their proportion of CpG motifs. In prokaryotic DNA, CpG motifs are present with a 20-fold excess as compared to eukaryotic DNA, which contains such motifs only temporarily, e.g. in cancer cells or promoter regions (Deutsches Ärzteblatt, Jg. 98/15: A981-A985 (2001)). In prokaryotic DNA, these motifs are not methylated, whereas most of them are methylated in eukaryotic DNA, which additionally increases their

35 distinctiveness. Non-methylated CpG motifs are non-methylated desoxycytidylate-desoxyguanylate dinucleotides in the prokaryotic genome or in fragments thereof.

The invention is further based on the finding that the protein according to the invention specifically binds to non-methylated CpG motifs. This specific binding property of the protein according to the invention is utilized in order to bind prokaryotic DNA and thus to subsequently enrich, separate and isolate it from a sample, e.g. with a majority of eukaryotic DNA.

The term "DNA containing non-methylated CpG motifs" refers to both eukaryotic and prokaryotic DNA. Said DNA can be purified and dissolved again (e.g. non-methylated DNA isolated from tissues) or be present directly in the original source (e.g. body fluid, such as blood, serum, tracheal aspirate, urine, bronchoalveolar lavage, nose smear, skin smear, puncture fluid).

According to a preferred embodiment, the DNA containing non-methylated CpG motifs is prokaryotic DNA, in particular bacterial DNA.

The term "homology" in the sense of the present invention relates to the degree of identity of two protein sequences. For example, a homology of 60 % means that 60 out of 100 amino acid positions in the sequences are identical. The term "shortened" used in order to characterize the protein according to the invention means that the length of the amino acid sequence of the protein according to the invention (e.g. CPGbP181) is shorter than the length of the amino acid sequence of the wild type CPGB protein (CPGbP656). Shortening is effected at the N-terminus and at the C-terminus of the wild type protein sequence (Figure 1). The maximum shortening is represented by the DNA binding site of the protein.

The protein employed according to the invention may have a molecular weight of, for example, approximately 19,959 Dalton (native) or 21,444 Dalton (in plasmid pQE60). In another preferred embodiment the isoelectric point of the protein according to the invention is approximately 10.09 (native protein) or 10.15 (in plasmid pQE60). A particularly preferred protein employed according to the invention has the amino acid sequence shown in SEQ ID No. 2 or in Fig. 1. This protein has particularly good binding properties as compared to non-methylated CpG motifs of prokaryotic DNA.

The protein described in EP 02020904 (CPGbP241), which is a shortened variant of the wild type CPGB protein (CPGbP656) and served as the basis for the protein employed according to the invention (e.g. CPGbP181), has a length of 241 amino acids, a molecular weight of approximately 33,650 Dalton (native) or 28,138 Dalton (in plasmid pQE60) and an isoelectric point of 9.89 (native) or 9.88 (in plasmid pQE60). The cDNA and amino acid sequence is shown in Figs. 1 and 2.

The wild type CGBP protein has a length of 656 amino acids, 135 positively charged residues and 94 negatively charged residues, a molecular weight of approximately 75,684 Dalton and an isoelectric point of 8.15. The cDNA and amino acid sequnce is shown in Fig.1.

5 The sequence comparison of the protein CPGbP181 according to SEQ ID No. 2 used according to the invention with the protein described in EP 01010904 (CPGbP241) is shown in Figs. 1 and 2.

The protein employed according to the invention is preferably produced by cloning the
10 corresponding cDNA sequence into a plasmid and by expression in *Escherichia coli*. An *E.coli* strain expressing the protein according to the invention was deposited with the Deutsche Sammlung für Mikroorganismen und Zellkulturen under No. DSM 16229 on February 16, 2004. Alternatively, other methods of manufacture known in the art can be applied. The use of plasmid pQE9 represents an exemplary possibility, but any other suitable plasmid is useful as a vector.
15 Expression in *E.coli* is also just an example. Expression in other prokaryotic systems and also in a eukaryotic system as well as chemical or enzymatic synthesis or purification from a natural source, such as e.g. tobacco plants, are further possible embodiments of protein extraction. The protein can be produced both on a laboratory scale (e.g. in an Erlenmeyer flask) and on an industrial scale (e.g. fermenter). For example, the protein according to the invention can be
20 purified by binding histidine residues (His-tag), which are introduced at the beginning or at the end of the protein, to a suitable nickel-containing matrix, which is a method known in the art.

Further possibilities of purification may be any type of fusion proteins allowing purification via suitable matrices (columns, gels, beads etc.). Other forms of tags may be fusion peptides/fusion
25 proteins, e.g. streptavidin-tag, Myc-tag and others.

A preferred form of the protein used according to the invention is the native form, but a denatured form is also suitable for binding non-methylated CpG motifs. "Denaturated forms" in the sense of the present invention are understood to be secondary structures other than those
30 found in nature.

The native or denaturated form of the protein used according to the invention is an exemplary embodiment. The invention includes *in vitro*-synthesis as well as any other chemical or enzymatic modifications of the protein, such as e.g. incorporation of disulfide bridges,
35 glycosilations, phosphorylations, acylations, amino acid exchanges as well as fusion with proteins or other molecules. Such modifications may be achieved, for example, by recombination and/or expression and/or chemical and/or enzymatic modification of single or multiple amino acids.

The protein used according to the invention has a multiplicity of advantages. It is better in binding prokaryotic DNA via non-methylated CpG motifs than the wild type CPGB protein or variants thereof with a homology of 80 % or more. This makes it possible to specifically separate and/or enrich the prokaryotic DNA of a mixture of prokaryotic and eukaroytic DNA. This ultimately enables quick and simple detection of pathogens as well as early diagnosis of infections which may be caused by bacterial pathogens. Conversely, the invention can also be used for depletion of microbial DNA in the sense of purification in the case of clinical conditions accompanied by non-physiological presence of bacteria or their cleavage products in body fluids, in particular blood, of patients. This applies even more because it is well documented that bacteria and also their cleavage products, such as, for example, bacterial DNA, are responsible for a multiplicity of biological effects detrimental to the patient.

Due to the good binding ability of the protein used according to the invention to non-methylated CpG motifs of prokaryotic DNA, the invention relates to a method of separating and/or enriching prokaryotic DNA, comprising the steps of:

a) contacting at least one prokaryotic DNA present in solution with a protein which specifically binds prokaryotic DNA and has 25 % to 35 % homology with the wild type CGPB protein, thus forming a protein-DNA complex, and

b) separation of said complex.

The DNA can be purified and dissolved again or may be present directly in the original source (e.g. body fluid, such as blood, serum, tracheal aspirate, urine, bronchoalveolar lavage, nose smear, skin smear, puncture fluid).

Separation may be effected by different methods of separating, isolating or enriching DNA protein complexes or DNA polypeptide complexes that are well-known to the person skilled in the art. In doing so, use will be made preferably of methods in which the DNA-binding protein is or is being immobilized to a carrier in order to separate and/or enrich the DNA from the sample solution.

According to a preferred embodiment, separation is followed by a step of separating the DNA from the protein according to the invention in said complex. This may be effected, for example, by conventional methods of DNA purification known to the person skilled in the art. In the most simple case, separation is effected by changing the pH value or the salt concentration (e.g. to 1 M NaCl) of the medium/buffer or by adding chaotropic reagents, etc.; i.e. suitable parameters

which lead to the separation of the protein-DNA-complex. Such methods are known to the person skilled in the art.

According to a further preferred embodiment, the protein according to the invention is coupled to a carrier. This embodiment represents a particularly simple way of enriching prokaryotic DNA, because separation from the solution is particularly easy, for example by means of physical removal (e.g. by centrifugation) of the charged carrier(s) from the solution.

For the solution of the prokaryotic DNA, any suitable solvent is basically contemplated. However, the method is particularly useful for enriching prokaryotic DNA from solutions which contain different biomolecular species, in particular different types of DNA. The invention preferably relates to a method of separating and enriching prokaryotic or viral DNA from a mixture of prokaryotic and eukaryotic DNA. In doing so, for example, the prokaryotic DNA which is present in body fluids is separated from the eukaryotic DNA by specific binding to the protein according to the invention and is enriched. The prokaryotic DNA enriched in this way facilitates detection of prokaryotic pathogens with the help of molecular biology methods and can contribute to the diagnosis of diseases caused by pathogenic pathogens.

In particular, the embodiment according to which the DNA-binding protein is immobilized to the surface of a carrier is suitable for adsorption of prokaryotic DNA from body fluids, preferably from blood. Moreover, this approach allows removal of microbial DNA, which is present in blood or other body fluids, from said fluids. The body fluid (e.g. whole blood, serum or liquor) purified in this way from the microbial DNA which is also capable by itself to initiate severe inflammatory reactions in patients, can then be fed back into the body. This principle may also be used for depletion of prokaryotic DNA from physiological fluids in the sense of purification, utilizing the specific binding properties of the protein according to the invention.

In order to increase the binding capacity and binding efficiency with respect to non-methylated CpG motifs of DNA, the invention provides a method enhancing the binding capacity and binding efficiency of the protein and thus allowing improved separation and/or enrichment of non-methylated DNA from a mixture of methylated DNA and non-methylated DNA.

According to the invention, this is achieved by indirect coupling of the protein to the matrix. This method will be described hereinafter with reference to Figures 9 and 10.

In order to enhance the binding capacity and binding efficiency of the CpGbP-181 protein with respect to DNA containing non-methylated CpG motifs, the invention relates to indirect binding of the protein to the matrix via a spacer. By coupling the protein to the matrix via a spacer, the

degree of mobility as well as the number of free binding sites of the CpGbP-181 protein is increased. Thus, increased binding capacity and binding efficiency are achieved. This further allows to reduce the amount of protein used.

5   Spacers in the sense of this invention are understood to be short chain-like molecules, which allow a spatial distance between the matrix and the protein used according to the invention, e.g. the CpGbP-181 protein. Such spacers are known in the art, e.g. from affinity chromatography or immobilization of proteins. Such chain-like molecules are composed of C and H atoms as well as optionally hetero atoms, e.g. N. These chain-like molecules are made up of individual chain

10   members on the basis of the C atoms, e.g. CH_2, and potentially present hetero atoms, e.g. NH. In particular, the spacer comprises 4 to 20, preferably 7 to 10 chain members. A particularly preferred spacer is derived from diamine hexane ($NH_2(CH_2)_6$-NH_2). Antibodies in the sense of the present invention are not to be considered as spacers.

15   A matrix in the sense of this invention relates to substances which function as carriers for the spacer and the protein. Carrier materials may be, for example, sepharose, pearl cellulose, silica, or similar substances known in the art.

      Body fluids in the sense of the invention are understood to be all fluids originating from the body

20   of a mammal, including humans, in particular such fluids in which disease pathogens may occur, such as blood, urine, liquor, pleural liquids, pericardial liquids, peritoneal liquids as well as synovial liquids. The description of the invention referring to human blood is not to be construed as limitative, but only as an exemplary application.

25   Bacterial pathogens are preferably understood to be pathogens of sepsis, but also any other bacterial pathogens of infections. They may differ from commensal pathogens, which are part of the normal population of the organism and are sometimes also found in test samples from patients, but do not have any clinical significance.

30   When isolating total DNA from infected body liquids, the ratio of host-DNA to pathogen-DNA may be, in many cases, only $1:10^{-6}$ to $1:10^{-8}$ or even less. Through the specific binding of prokaryotic DNA to the protein according to the invention, the method according to the invention enables enrichment by 1 exponential unit and more.

35   The protein used according to the invention may be coupled directly or indirectly to the carrier. The type of coupling depends on the carrier and the carrier material. Suitable carriers include, in particular, membranes, microparticles and resins, or similar materials for affinity matrices. Suitable materials for binding the protein according to the invention, as well as – depending on

the type of material – for carrying out such binding are well-known to the person skilled in the art. For indirect coupling, specific antibodies against the protein according to the invention or the polypeptide are suitable, for example, which are in turn bound to the carrier by known methods.

5   One application of the method according to the invention consists in enriching prokaryotic DNA. A further application consists in the separation of prokaryotic DNA from a mixture of eukaryotic and prokaryotic DNA by binding the prokaryotic DNA to the protein used according to the invention, which has been immobilized, for example, to a matrix. The mixture of the body's own DNA and prokaryotic DNA is contacted with the affinity matrix by means of suitable methods

10   and, in doing so, the prokaryotic DNA is bound to the immobilized protein; the eukaryotic DNA passes, for example, through a separating column and may be collected separately. Affinity matrices may be, for example, polymeric polysaccharides, such as agaroses, other biopolymers, synthetic polymers, or carriers having a silicate backbone, such as porous glasses or other solid or flexible carriers on which the DNA-binding protein used according to the

15   invention is immobilized. After separation of prokaryotic DNA from eukaryotic DNA has been effected, the affinity matrix is rinsed with a suitable reagent, so that the binding protein with the coupled prokaryotic DNA is separated from the matrix and/or the prokaryotic DNA is separated from the binding protein and is available in a sufficient amount for further process steps.

20   A further application of the method according to the invention consists in the separation and enrichment of prokaryotic DNA from eukaryotic DNA by binding the prokaryotic DNA to the protein according to the invention, which has been immobilized on microparticles. In this connection, all microparticles which allow the DNA-binding protein according to the invention to be immobilized are suitable. Such microparticles may consist of latex, plastics (e.g. styrofoam,

25   polymer), metal, or ferromagnetic substances. Furthermore, use may also be made of fluorescent microparticles, such as those available from the Luminex company for example. After the prokaryotic DNA has been bound to the proteins used according to the invention, which are immobilized on microparticles, said microparticles are separated from the mixture of substances by suitable methods, such as filtration, centrifugation, precipitation, sorting by

30   measuring the intensity of fluorescence, or by magnetic methods. After separation from the microparticles, the prokaryotic DNA is available for further processing.

Another application of the method according to the invention consists in the separation and enrichment of prokaryotic DNA from eukaryotic DNA by binding the prokaryotic DNA to the

35   protein used according to the invention, which is subsequently separated from other ingredients of the mixture by electrophoresis.

A further application of the method according to the invention consists in the separation and enrichment of prokaryotic DNA from eukaryotic DNA by binding the prokaryotic DNA to the protein used according to the invention. The protein used according to the invention is subsequently bound to corresponding antibodies. The antibodies may be bound to solid or

5   flexible substrates, such as glass, plastics, silicone, microparticles, membranes, or may be present in solution. After binding of the prokaryotic DNA to the protein according to the invention and binding of the latter to the specific antibody, separation from the substance mixture is effected by methods known to the person skilled in the art.

10   The method according to the invention may also be used in order to purify body fluids to remove prokaryotic DNA therefrom. In this connection, it is convenient for separation to be effected extra corporally, under sterile conditions, to allow the body fluids to be fed back into the body again, so that the body's own immune system is assisted in eliminating infections by removing the prokaryotic DNA contained in said body fluid.

15

Any suitable chemical, mechanical or electrochemical processes may be considered for extracorporal removal of prokaryotic DNA from body fluids. Further, the combination with other extracorporal methods, such as hemoperfusion, heart-lung machine or endotoxin adsorbers, is a further convenient application.

20

The protein used according to the invention can also be used to detect prokaryotic DNA. In this case, enrichment of the prokaryotic DNA is followed by a step of amplifying said prokaryotic DNA, for which all common methods of amplification are suitable (PCR, LCR; LM-PCR , etc.).

25   The method according to the invention, in particular with the above-described embodiments, has the advantage that, by specific binding of non-methylated prokaryotic DNA, rich in CpG motifs to proteins with specific affinity for such structures, prokaryotic DNA from the total DNA of an infected host is successfully concentrated and thus the sensitivity of detection of pathogen DNA in body fluids is strongly enhanced.

30

The possibilities of separating prokaryotic DNA from eukaryotic DNA using a specifically binding protein are no more time-consuming than known methods of isolating total DNA. However, subsequent detection can then be effected only by PCR. A nested PCR will not be required in most cases, which makes it possible to save a considerable amount of time in diagnostics.

35

The use of the protein of the invention to deplete prokaryotic DNA in physiological body fluids was already mentioned above. Depletion in the sense of the present invention means that the amount of prokaryotic DNA is reduced. This possibility of reducing prokaryotic DNA also

enables the use of the proteins according to the invention in environmental technology, waste water management and air conditioning technology.

The invention further relates to a method of separating and enriching non-methylated genomic DNA from a mixture of non-methylated genomic and methylated genomic DNA. The methylated genomic DNA is separated by binding the non-methylated genomic DNA to the CpGbP-181 protein coupled to a matrix. This procedure contributes substantially to the simplified examination of the methylation patterns of methylated genomic DNA and enables the diagnosis of diseases having a specific methylation pattern.

Moreover, the invention relates to a kit for enriching prokaryotic DNA by one of the above-described methods, said kit containing at least the protein according to the invention, optionally together with further reagents suitable to carry out said method.

In addition to the protein according to the invention, said kit may contain at least one set of primers, which are suitable to amplify genomic DNA of certain prokaryonts under standard conditions.

The invention will be explained in more detail below with reference to the examples, without limiting it thereto.

Example 1: Preparation of the protein according to the invention

The DNA sequence for the complete CPGbP protein was used to construct primers 1 (GGATCCGGTGGAGGGCGCAAGAGGCCTG –fw SEQ ID No. 3) and 2 (AAGCTTAGAGGTAGGTCCTCAT-CTGAG-rv SEQ ID No. 4) which amplify a shortened DNA fragment encoding CPGbP-181, which is a shortened protein binding CPG. After cleavage, the DNA fragment was ligated into the pQE9 vector (Qiagen) using restriction enzymes *BamHI* and *Hind III*. An open reading frame forms in pQE9, in which frame a DNA fragment encoding 6 x His-Tag (pQE9[6HisCPGbP181]) is fused to the 5' end. The complete amino acid sequence of the encoding fusion protein 6His-CPGbP181 is shown hereinafter, the portion indicated in bold print representing the peptide CPGbP181 and the portions indicated in italics indicating fused foreign amino acids of plasmid pQE9.

Plasmid pQE9[6HisCPGbP181] was transformed to the *E. coli* expression strain M15[pREP4] (Qiagen). The clone is referred to hereinafter as M15[pCPGbP181], and the expressed protein is referred to as rCPGbP181. Expression of the protein rCPGbP181 occurred according to the following protocol: A colony of the expression strain M15[pCPGbP181] is grown overnight in 2

ml Luria Medium with 100 µg/ml ampicilline and 25 µg/ml kanamycine at 37°C with shaking. Then, the pre-culture is transferred to 200 ml preheated nutrient medium containing the same concentrations of antibiotics. After 3 hours of growth at 37°C with shaking, IPTG is added to induce expression, and incubation is continued for 5 hours. Thereafter, the bacteria are

5 removed by centrifugation and the sediment is re-suspended in 5 ml 0.2 M tris buffer, pH 7.5. The bacteria are subjected to ultrasonic treatment in an iced bath for 5 x 1 min. After centrifugation, the sediment is re-suspended in 10 ml 0.2 M tris, 2M urea, pH 7.5, and shaken for 15 min. After centrifugation has been effected, the remaining sediment is taken up in 0.2 M tris, 6M guanidine hydrochloride, 0.001 M dithioeritrite (DTE), 0.02 M imidazole, and suspended

10 therein. The inclusion bodies are dissolved at room temperature for 1 hour with agitation. After centrifugation, the crude protein is present in the supernatant and can be applied directly to a 3 ml Ni-agarose column. The subsequent steps should be effected in the cooling chamber at +4 to +6°C. First, the column is washed with 0.2 M tris, 6M guanidine hydrochloride, 0.001 M dithioeritrite (DTE), 0.02 M imidazole buffer, pH 7.5, until extinction has reached the zero line.

15 From this point, rCPGbP181 can be obtained in different ways: 1. as a denatured protein, dissolved in 6M guanidine hydrochloride or 6M urea, and 2. as a native protein, soluble in buffers at physiological concentrations. In the second case, however, the yield is lower.

Purification according to method 1 (denaturated):

20 The protein rCPGbP181 is eluted from Ni-NTA agarose with an imidazole gradient of 0 – 0.5 M, M in buffer 0.2 M tris, 6M guanidine hydrochloride, 0.001 M dithioeritrite (DTE), 0.02 M imidazole, pH 7.5, as the basic material. In doing so, rCPGbP181 is detached from the column at 0.2 – 0.3 M imidazole. The protein thus obtained is dialyzed against 0.2 M tris, 6M urea, 0.001 M dithioeritrite (DTE), pH 7.5, and frozen. During dialysis against physiological buffers,

25 purified rCPGbP181 is thus precipitated.

Purification according to method 2 (native):

According to this method, the guanidine hydrochloride concentration is shifted from 6 mol on Ni-NTA agarose with the bound rCPGbP181 via a gradient up to 0 mol guanidine hydrochloride.

30 The basis for this is the buffer 0.2 M tris, 0.5 M NaCl, 0.001 M dithioeritrite (DTE), 0.02 M imidazole, pH 7.5. In this case, a flow rate of 0.5 ml/min was selected. Subsequently, an imidazole gradient of 0 to 0.5 mol was applied for elution in buffer 0.2 M tris, 0.5 M NaCl, 0.001 M dithioeritrite (DTE), pH 7.5, as basic material. In this case, too, a substantial proportion of the bound protein (20 %) was eluted at 0.2 to 0.3 mol imidazole. This native rCPGbP181 eluate

35 remained dissolved in this buffer even after dialysis in PBS. However, it is disadvantageous that approximately 80 % of rCPGbP181 bound to Ni-NTA agarose remained on the column under these conditions and were subsequently extractable only under the denaturated conditions of

method 1. This means that the yield of method 2 as used resulted only in 20 % native rCPGbP181 soluble in physiological buffers.

Example 2. Detection of pathogens by means of nested PCR.

5   Fresh, heparinized human blood, which contains *streptococcus pyogenes* with 10^3/ml colony-forming units as pathogens, is used for detection of pathogens. The DNA is isolated by means of absorption to DNA-binding matrix using commercial kits for isolation of total DNA from body fluids according to modified instructions from the manufacturers. For this purpose, 200 μl of the total lysis buffer, which contains proteinase K and SDS, is added to

10   100 μl of infected blood in Eppendorf tubes. The mixture is incubated at 37°C for 30 min and then heated to 95°C for 20 min. After cooling, 20 μg of mutanolysine are added and incubated at 37°C for another 60 min. After centrifugation, the mixture is applied to the centrifugation columns using DNA-binding matrix and the DNA is purified according to manufacturer's instructions. The purified DNA is placed in a final volume of 100μl of 0.01 mol

15   tris buffer, pH 7.5, or in an equal amount of elution buffer from the manufacturer. For detection of pathogens, primers were selected to identify the streptolysin O gene (slo).

1. PCR.   Amplification of a 465 bp fragment

Forward primer 1: 5'-AGCATACAAGCAAATTTTTTTACACCG

20   Reverse primer 2: 5'-GTTCTGTTATTGACACCCGCAATT

Primer concentration 1mg/ml

Starting material: 5 μl isolated DNA

            0.5 μl primer fw 1

            0.5 μl primer rv 2

25   14 μl aqua dest

            total 25 μl in Ready to go Kit (Amersham-Pharmacia)

Reaction:

5 min 95 °C

30   40 cycles (30 sec. 95°C; 30 sec. 51°C; 3 min 72°C; 1 x 7 min 72°C).

The first PCR of streptococci-DNA in human blood is shown in Fig. 1 (10 μl each of the 25 μl starting material were separated. 1) PCR starting material containing 5 μl template DNA; 2) starting material containing 5 μl template, at a dilution of 1:10. 3) positive control: 0.2 μl of

35   streptococci-DNA as template in the absence of eucaryotic DNA from blood. ST) molecular weight standard)

Result: The first primary PCR does not result in a positive reaction. Therefore, a second PCR (nested PCR) was subsequently carried out.

2. PCR (nested): Amplification of a 348 bp fragment in the above *slo*-fragment.

Forward primer 3: 5'-CCTTCCTAATAATCCTGCGGATGT

Reverse primer 4: 5'-CTGAAGGTAGCATTAG TCTTTGATAACG

Primer concentration: 1mg/ml

Starting material: 5 µl from PCR1, sample 1, Fig. 1

      0.5 µl primer fw 1

      0.5 µl primer rv 2

      14 µl aqua dest

      total 25 µl in Ready to go Kit (Amersham-Pharmacia)

Reaction:

5 min 95°C

40 cycles (30 sec. 95°C; 30 sec. 54°C; 3 min 72°C; 1 x 7 min 72°C)

Figure 4 shows the nested PCR with the PCR products from the primary PCR starting material according to Fig. 3 as template. The samples correspond to those of Figure 3.

Result: In the nested PCR, the desired slo-DNA fragment is amplified at a concentration of 100 streptococci cells per 100 µl blood (sample 1). For 5 µl starting material in the 1st PCR (Fig. 3), this corresponds to about 5 to 10 templates. At a dilution of 1:10 (sample 2), sensitivity is exhausted (0.5 to 1 template).

These experiments show that successful PCR detection of pathogens in blood requires isolation of the total DNA from at least 1 to 5 ml blood. However, the total DNA concentration is then too large to be used directly in a PCR.

Other pathogen-specific nucleic acid detections without an amplification step by direct detection of the bacterial DNA, for example by DNA hybridization, are also too insensitive, which is primarily due to the high excess of human DNA relative to bacterial DNA. In addition, competitive processes during DNA analysis as well as the low quantity of bacterial DNA are to be regarded as hindrances to qualitative and quantitative analysis. The common methods of DNA isolation enrich the total DNA of a body fluid so that the ratio of host DNA to microbial DNA can be between $1:10^{-6}$ and $1:10^{-8}$. This difference makes it easy to understand the difficulty in detecting microbial DNA in body fluids.

Example 3: Determining the binding properties of rCPGbP181:

In gel retardation experiments both the binding of the denaturated and of the native protein rCpGbP181 to methylated and to non-methylated DNA molecules with CpG motifs was examined. The pUC18 plasmid of *E. coli* was used as the test DNA with an inserted M-protein gene segment of *streptococcus dysgalactiae supsp. equisimilis* (Geyer et. al FEMS Immuno. Med. Microbiol. 26:11-24, 1999). The plasmid preparation was divided and one half was methylated with the CpG methylase kit of New England BioLabs. Both preparations were mixed with rCPGbP181 (native or denaturated) and electrophoretically separated on agarose gel. The results are shown in Figures 5 and 6. Both the native form and the denaturated form of rCPGbP181 showed a higher affinity to non-methylated plasmid DNA, which confirms the selective binding property with respect to non-methylated CpG-rich DNA.

Description of the gel retardation experiment according to Figure 5: 5µl (72 ng) methylated pUC18*emm* DNA and 1 µl (142 ng) non-methylated pUC18*emm* DNA, respectively, were mixed with 5 µl (0.5 µg) native rCPGbP181 and filled up to a volume of 35 µl with the following buffer: 0.01 M tris, 0.08M NaCl, 0.001M EDTA, 0.005M DTE, 5% glycerine, pH 7.8. After incubation at 20°C for 30 min the mixtures were electrophoretically separated on 1.5% agarose. Methylated DNA was applied in lanes 1 and 3 and non-methylated DNA was applied in lanes 2 and 4. In lanes 1 and 2 the DNA was mixed with native rCPGbP181. Lane 2 shows that non-methylated pUC18*emm* interacts with rCPGbp181; in contrast thereto, rCPGbP181 did not show any interaction with methylated pUC18*emm* (lane 1). Lanes 4 and 5 are the plasmids without addition of rCPGbP181 as controls.

Description of the gel retardation experiment of Figure 6 for non-methylated and methylated pUC18*emm* after incubation with denaturated rCPGbP181. The concentrations correspond to those of Figure 5. Methylated DNA was applied in lanes 1 and 3 and non-methylated DNA was applied in lanes 2 and 4. In lanes 1 to 4, the DNA was mixed with two different batches of denaturated rCPGbP181. Lanes 2 and 4 show that non-methylated pUC18*emm* also interact with denaturated rCPGbP181; however, rCPGbP181 did not show any interaction with methylated pUC18*emm* (lanes 1 and 3). Lane 5 is pUC18*emm* without rCPGbP181 as control.

Example 4: Binding and separation of a mixture of calf thymus DNA and bacterial DNA to immobilized CPGbp181.

Purified CPGbp181 was coupled to aminohexyl sepharose (Amersham-Biosciences) by means of glutaraldehyde according to the protocol of Cambiasso et al. (Cambiasso, C. et al., Immunochemistry 12-273-278, 1975). The concentration of immobilized protein was 0.3 mg per

milliliter sepharose. 300 µl sepharose was placed in a spin-filter tube containing inert fritting material which absorbs neither DNA nor protein, but retains sepharose.

200 ng calf thymus DNA and 25 ng pUC18emm was dissolved in 100 µl 20 mM tris-HCL buffer,
5   pH 7.5, and applied to the column thus prepared. After each step, the liquid was centrifuged at 14,000 RPM for 0.5 min in an Eppendorf centrifuge in one fresh Eppendorf tube each. Thus, a two-step increase of the NaCl concentration was effected from 0 to 1M. DNA precipitation was effected in each tube by adding 10 µl 4 M acetate, ph 4.5, and 250 µl ethanol abs., mixing and centrifugation at 14,000 RPM for 15 min. Thereafter, the supernatant was discarded and the
10 precipitate was washed with 300 µl 70 % ethanol. After discarding, the residue was dried for 5 min in a vacuum centrifuge and then taken up in 15 l distilled water (PCR-suitable). On the one hand, extinction at 254 nm was measured for 10 µl each of the samples (Figure 7). On the other hand, PCR was effected with sequence primers for PUC18, using 3 µl of each sample (Figure 8).

15

The result (Figures 7, 8) shows that the eukaryotic calf thymus DNA is initially washed from the column between 0 to 0.1 M NaCl, while the prokaryotic DNA (pUC18emm) was eluted in the fraction at 0.3 M NaCl. This shows that eukaryotic DNA has a lower affinity to CPGbP181 and, thus, a clear separation of both DNA fractions was achieved.

20

Example 5: Enhancement of binding properties of the CpG-bP-181 protein, which result from indirect binding of this protein to a matrix via a spacer.

In order to examine binding properties, prokaryotic DNA from a DNA mixture of *staphylococcus*
25 *aureus* and human DNA was enriched using the directly coupled CpGbP-181 protein on CNBr sepharose or using the indirectly coupled CpGbP-181 protein on sepharose (in the following AH sepharose) via a diaminohexyl spacer (AH).

First, the AH sepharose was incubated at room temperature for 15 min with addition of
30 glutaraldehyde. Next, the AH sepharose was washed with 0.1 mol Na_2HPO_4. Then, 0.24 mg of the CpGbP-181 protein was placed on the matrix. The binding of the CpGbP-181 protein to AH sepharose was achieved by incubation at room temperature for 2 hours. The excess CpGbP-181 protein was removed.

35 After subsequent washing of the CpGbP-181-AH sepharose with 0.1 mol Na_2HPO_4 and addition of 0.1 mol glycine, the CpGbP-181-AH sepharose was incubated at room temperature for 2 hours in order to saturate free binding sites. Then, the CpGbP-181-AH sepharose was again washed with 0.1 mol Na_2HPO_4. In order to reduce the Schiff base and in order to stabilize

binding, the CpGbP-181-AH sepharose was admixed with sodium borohydride and incubated at room temperature for 1 hour.

Next, the CpGbP-181-AH sepharose was washed with 0.1 mol Na_2HPO_4.

5

The storability of the CpGbP-181-AH sepharose at 4°C is achieved by addition of 20% ethanol. Next, the CpGbP-181-AH sepharose was portioned into columns. The columns prepared with CpGbP-181-AH sepharose were then washed with tris buffer and were available for separation/enrichment of DNA containing non-methylated CpG motifs.

10

2) Enrichment of the DNA mixture, followed by elution of the prokaryotic DNA and determination of the concentration of prokaryotic DNA by PCR.

The DNA mixture consisted respectively of 330 ng human DNA and 150 ng prokaryotic DNA
15 (*staphylococcus aureus* DNA). The DNA mixture was placed on the columns prepared with CNBr sepharose or with AH sepharose, respectively, and incubated at room temperature for 1 minute. Then, the columns were centrifuged and washed with 100 l tris buffer (10 µM, pH 7). The washing and centrifugation step was repeated 5 times.

20 The supernatant was carefully removed and then 100 µl elution buffer (10 µM tris buffer, 0.5 M NaCl, pH 7) each were added to the columns and centrifuged. The elution step was repeated 5 times. Then, the individual fractions of each sample were precipitated by addition of 10 µl 3 M sodium acetate and 250 µl ethanol with subsequent mixing and centrifugation (15 min at 15,000 g). The supernatant was carefully discarded and the pellet was washed with 1 µl ethanol (70%)
25 and centrifuged at 15,000 g. The supernatant was then removed again, the pellet was dried in a vacuum centrifuge and taken up in 30 µl DEPC water. 5 µl each were used for PCR detection.

The PCR used universal primers for the 16s RNA gene. After carrying out the PCR, 15 µl each of the individual fractions were placed on 2% agarose gel.

30

Figure 9 (direct binding of the CpG-181 protein to CNBr sepharose) and Figure 10 (indirect binding of the CpG-181 protein to sepharose via a spacer (AH)) show the results of the PCR for the individual fractions. It is clearly evident that the use of the AH spacer allowed more prokaryotic DNA to be enriched (fraction 1, elution fraction). This characteristic improvement in
35 binding properties is useful in the methods according to the invention.